Energy-Aware Wireless Networking with Directional Antennas: The Case of Session-Based Broadcasting and Multicasting

Jeffrey E. Wieselthier, *Senior Member*, *IEEE*, Gam D. Nguyen, and Anthony Ephremides, *Fellow*, *IEEE*

Abstract—We consider ad hoc wireless networks that use directional antennas and have limited energy resources. To explore quantitatively the advantage offered by the use of directional antennas over the case of omnidirectional antennas, we consider the case of connection-oriented multicast traffic. Building upon our prior work on multicasting algorithms, we introduce two protocols that exploit the use of directional antennas and evaluate their performance. We observe significant improvement with respect to the omnidirectional case, in terms of both energy efficiency and network lifetime. Additionally, we show that further substantial increase in the network's lifetime can be achieved by incorporating a simple measure of a node's residual energy into the node's cost function.

Index Terms—Broadcast, multicast, energy efficient, directional antenna, ad hoc network.

1 Introduction

The use of directional antennas can provide energy savings and interference reduction by concentrating RF energy where it is needed. Hence, they are especially useful in networks with finite energy resources and a limited number of available frequencies. In this paper, we develop and evaluate algorithms for broadcasting and multicasting that are suitable for use in networks with directional antennas and limited battery capability, and compare performance to that achieved when antennas are omnidirectional. We focus on the problem of tree construction for source-initiated, session-based traffic in all-wireless (i.e., infrastructureless, peer-to-peer, or ad hoc) multihop networks.

In our earlier studies [1], [2], we developed energy-aware algorithms for the construction of broadcast and multicast trees for networks with omnidirectional antennas. In the broadcasting case, the goal is to find a minimum-cost (e.g., minimum power) tree, rooted at the Source node, that reaches all other nodes in the network; in the multicasting case, a specified subset of the nodes must be reached, while others may serve as relays. These algorithms are known as Broadcast Incremental Power (BIP) and Multicast Incremental Power (MIP) [1], [2], respectively.[1] It has subsequently been proven that the minimum-cost broadcasting problem in wireless networks with omnidirectional antennas is NP-complete [7], [4]. By contrast, the minimum-cost broadcasting problem in wired networks can be formulated as the well-known, and easily solved, minimum-cost spanning tree (MST) problem.

BIP and MIP are examples of "node-based" algorithms, which exploit the "wireless multicast advantage" property associated with omnidirectional antennas, namely, the capability for a node to reach several neighbors by using a transmission power level sufficient to reach the most distant one. This property no longer holds for directional antennas because the RF power needed to sustain sufficient signal levels depends on the beamwidth. In this paper, using the incremental power philosophy as a starting point, we demonstrate the issues that arise when directional antennas are used, we develop algorithms for tree construction, and we illustrate trade-offs between complexity and performance.

We address energy-aware networking from the perspectives of both energy-efficient and energy-limited operation. In the former case, it is implicitly assumed that energy reserves (batteries) can be renewed during the course of network operation, and a cost is associated with the quantity of energy that is expended. In the latter case, each node is equipped with batteries that cannot be recharged during network operation. Thus, there is a hard constraint of a *fixed quantity of energy at each of the network nodes*. We briefly address some of the fundamental differences between energy-limited and energy-efficient network operation. Our primary focus in this paper is on energy-limited operation.

In [8], for the case of omnidirectional antennas, we studied the performance of MIP under the constraint of a fixed quantity of energy at each of the network nodes. We demonstrated that the lifetime of the network can be

1. BIP has also been studied by other researchers [3], [4], and [5]. Alternative algorithms, which perform better than BIP/MIP but at the expense of higher complexity, were introduced in [6].

- J.E. Wieselthier and G.D. Nguyen are with the Information Technology Division, Naval Research Laboratory, Code 5521, Washington, DC 20375. E-mail: {wieselthier, nguyen}@itd.nrl.navy.mil.
- A. Ephremides is with the Electrical and Computer Engineering Department and the Institute for Systems Research, University of Maryland, College Park, MD 20742. E-mail: tony@eng.umd.edu.

Manuscript received 14 June 2002; accepted 18 July 2002.
For information on obtaining reprints of this article, please send e-mail to: tmc@computer.org, and reference IEEECS Log Number 9-042002.

1536-1233/02/$17.00 © 2002 IEEE

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE **2002**	2. REPORT TYPE	3. DATES COVERED **00-00-2002 to 00-00-2002**
4. TITLE AND SUBTITLE **Energy-Aware Wireless Networking with Directional Antennas: The Case of Session-Based Braodcasting and Multicasting**		5a. CONTRACT NUMBER 5b. GRANT NUMBER 5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)		5d. PROJECT NUMBER 5e. TASK NUMBER 5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Naval Research Laboratory,Information Technology Division,4555 Overlook Avenue, SW,Washington,DC,20375**		8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)		10. SPONSOR/MONITOR’S ACRONYM(S) 11. SPONSOR/MONITOR’S REPORT NUMBER(S)
12. DISTRIBUTION/AVAILABILITY STATEMENT **Approved for public release; distribution unlimited**		
13. SUPPLEMENTARY NOTES		
14. ABSTRACT		
15. SUBJECT TERMS		

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**		**16**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

extended significantly by incorporating into the tree-construction process a cost-function that reflects the residual energy at the nodes. Here, we demonstrate that, when modified to exploit the properties of directional antennas, the incremental power philosophy of BIP and MIP can provide further substantial gain in both energy efficiency and network lifetime.

Few studies of the use of directional antennas for energy-aware operation have appeared in the literature. Our preliminary results on the work presented in this paper appeared in [9], and the complete model was presented in [10]. The use of directional antennas to provide improved energy efficiency and interference reduction in multicast applications (but assuming that the trees are already known) was later studied in [11]. Other applications of the use of directional antennas for energy efficiency were studied in [12].

The focus of this paper is the construction of energy-aware broadcast or multicast trees, specifically for the case of directional antennas. To provide a better understanding of the fundamental issues involved, as in [1], [2], [8], we do not address many of the implementation issues that will have to be addressed eventually. For example, we do not address admission control; we assume that a tree is constructed to reach all destinations that are, in fact, reachable, regardless of the cost (energy) to do so. Nor do we address the control of transmission rate or other transmission parameter (which can affect session duration, energy usage, quality of service, etc.). We also choose to assume that the channel bandwidth and signal design parameters are set so that the bit rate is fixed. Additionally, we do not address media-access issues or mobility. Thus, our study of tree construction in the wireless environment translates to choosing transmission power and the set of receiving neighbor nodes at each level in the multicast tree.

In Section 2, we provide a brief discussion of the differences between energy-constrained and energy-efficient communication. In Section 3, we discuss the communication model, including the impact of the use of directional antennas on energy consumption. In Section 4, we define the multicasting problem, including the performance measures used in this study. In Section 5, we present our algorithms for broadcast and multicast-tree construction, which exploit the properties of directional antennas. In Section 6, we show how the impact of resource limitations (transceivers, frequencies, and energy) can be incorporated into our model. In Section 7, we present our performance results for energy-constrained systems and demonstrate that significant increase in delivered traffic volume can be achieved by using algorithms that exploit the properties of directional antennas. Finally, in Section 8, we present our conclusions from this research.

2 Energy-Constrained vs. Energy-Efficient Communication

The introduction of hard constraints on the total amount of energy available at each node results in a problem that is very different from that in which unlimited energy is available (although energy efficiency still may be desired). Under such hard constraints on energy, the network is capable of operation for a limited period of time. A node dies (and, hence, can no longer transmit) when its energy is depleted, and the network dies when it is no longer capable of providing a minimum acceptable level of service. By contrast, when the goal is energy efficiency (e.g., delivering the largest number of bits per unit energy), it is implicitly assumed that ample energy is available; in such cases, the use of energy is essentially treated as a cost function.

Energy-efficient operation does not ensure good performance in *energy-constrained* applications. For example, use of the most energy-efficient routes (or multicast trees) may result in premature depletion of energy at some nodes. We use the term *energy-aware* to refer to systems designed with either of these criteria in mind.

A more complete discussion of energy-limited versus energy-efficient communication is provided in [8], where performance results are provided for networks that use omnidirectional antennas.

3 The Model

We consider source-initiated, circuit-switched, multicast sessions. The maintenance of a session requires the dedication of a transceiver at each participating node (source node, relay nodes, and destination nodes), as well as the needed amount of interference-free bandwidth, throughout the duration of the session. The network consists of N nodes, which are randomly distributed over a specified region. Each node has T transceivers and can thus support up to T multicast sessions simultaneously. We assume that there is a total of F frequencies available to the network. Frequencies can be reused, provided that doing so does not create interference. Thus, congestion (and, hence, the inability to reach one or more destinations) may arise when either an insufficient number of transceivers or an insufficient number of frequencies are available at one or more nodes along the path. Alternatively, energy-inefficient paths may have to be used when the best paths are not available.

It is also of interest to study systems that use time-division multiple access (TDMA), rather than multiple transceivers or multiple channels, to support multiple sessions simultaneously. In TDMA-based systems, the need to assign specific time slots creates a much more difficult problem than that of simply assigning any transceiver (of perhaps several available) to a new session. Alternatively, it would be possible to consider code-division multiple access (CDMA) [13]. The study of TDMA- and CDMA-based systems is not pursued here since we want to place emphasis on the energy constraint with as little complication from the MAC layer as possible.

Any node is permitted to initiate multicast sessions. Multicast requests and session durations are generated randomly at the network nodes. Each multicast group consists of the source node plus at least one destination node. Additional nodes may be used as relays either to provide connectivity to all members of the multicast group or to reduce overall energy consumption. The set of nodes that support a multicast session (the source node, all destination nodes, and all relay nodes) is referred to as a

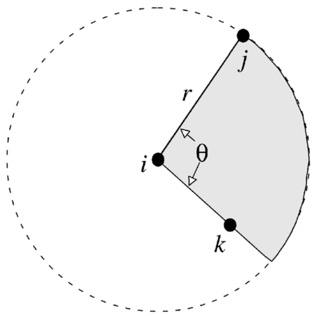


Fig. 1. Use of directional antenna to reach two neighboring nodes.

multicast tree. Notice the difference between this definition and the conventional one that is based on links (or edges); here, the links are incidental and their existence depends on the transmission power of each node. Thus, it is the nodes (rather than the links) that are the fundamental units in constructing the tree.

The connectivity of the network depends on the transmission power and antenna pattern. We assume that each node can choose its RF power level p^{RF}, such that $p_{\min} \leq p^{RF} \leq p_{\max}$. The nodes in any particular multicast tree do not necessarily have to use the same power levels; moreover, a node may use different power levels for the various multicast trees in which it participates.

3.1 Propagation Model

When considering omnidirectional antennas and uniform propagation conditions, we assume that the received signal power is equal to $pr^{-\alpha}$, where p is the transmission power, r is the distance, and α is a parameter that typically takes on a value between two and four, depending on the characteristics of the communication medium. Based on this model, the transmitted power required to support a link between two nodes separated by distance r is proportional to r^{α} since the received power must exceed some threshold.[2] Let us first consider the case of point-to-point communication between Node i and Node j. Without loss of generality, we set the proportionality constant equal to 1, resulting in:

$$\begin{aligned} p_{ij}^{RF} &= \text{RF power needed for link between Nodes } i \text{ and } j \\ &= \max\{r_{ij}^{\alpha}, p_{min}\}, \end{aligned} \tag{1}$$

where r_{ij} is the distance between Node i and Node j. The use of a nonzero value of $p_{\min}$ is a way to account for the fact that the $r^{-\alpha}$ dependence applies only in the far-field region (i.e., even when two nodes are arbitrarily close to each other, a nonzero power level $p_{\min}$ is required to support communication between them). As a consequence of the "wireless multicast advantage" property of omnidirectional systems [14], all nodes whose distance from Node i does not exceed r_{ij} will be able to receive the transmission with no further energy expenditure at Node i.

2. This threshold depends on factors such as signal parameters, detector structure, and noise levels (including other-user interference). In this paper, we assume that these characteristics are fixed; thus, the required level of received power is the same at all nodes. Thus, we neglect fading effects that arise in wireless channels.

The use of directional antennas can permit energy savings and reduce interference by concentrating transmission energy where it is needed. On the other hand, since only the nodes located within the transmitting node's antenna beam can receive the signal, the effect of the wireless multicast advantage may be diminished; additional power is needed as the beamwidth is increased to include additional receiving nodes. We use an idealized model in which we assume that all of the transmitted energy is concentrated uniformly in a beam of width θ, as shown in Fig. 1, where Node i transmits to Node j and Node k.[3] Then, the RF power needed by a node to transmit to a distance r using beamwidth θ (and thus to reach all nodes in this sector) is

$$p^{RF}(r, \theta) = \max\left\{ r^{\alpha} \frac{\theta}{360}, p_{\min} \right\}. \tag{2}$$

Consequently, the use of narrow beams permits energy saving (for a given communication range) or range extension (for a given transmitter power level), as compared to the use of omnidirectional antennas. Specifically, for a given value of $p_{\max}$, the maximum range is increased by a factor of $(360/\theta)^{1/\alpha}$, compared with the case of omnidirectional antennas.

We assume that the beamwidth θ can be chosen so that $\theta_{\min} \leq \theta \leq \theta_{\max}$. Furthermore, we assume that each node knows the precise locations of its potential neighbors (i.e., all nodes that can be reached with acceptable signal-to-interference ratio when transmitted with $p^{RF} \leq p_{\max}$) and that each antenna beam can be pointed in any desired direction to provide connectivity to a subset of the nodes that are within communication range. (In practice, the number of antenna elements needed tends to increase as $\theta_{\min}$ decreases.)

We assume that one antenna beam can be supported for each session in which a node participates; thus, the use of directional antennas does not have an impact on the number of sessions that a node can support simultaneously (as compared to an implementation with omnidirectional antennas). Additionally, both θ and the direction in which the beam points are chosen independently for each session in which a node participates. Although setting $\theta = \theta_{\min}$ is appropriate for point-to-point applications, it is often appropriate to use larger values of θ in multicast applications since a node may have several downstream neighbors, all of which must be included in a single beam (based on the assumption just made above). We discuss the choice of θ in our discussion of multicast algorithms in Section 5.

Although we do consider energy expenditures associated with processing at each node (in addition to that for RF transmission), we do not explicitly connect the amount of processing energy with the beamwidth of the antenna. This coupling is deferred for future investigation.

The use of directional receiving antennas would also have a beneficial impact since background noise and other-user interference would be troublesome only when located within the antenna beamwidth rather than the entire omnidirectional region. Thus, lower signal levels would be needed to provide the required performance. However,

3. In realistic systems, antenna patterns would be characterized by nonuniform gain in the primary sector of coverage, as well as by sidelobes.

we assume the use of omnidirectional receiving antennas to simplify the model.

It is also possible to consider alternative models, which may incorporate one or more of the following:

- fixed beamwidth (i.e., $\theta_{\min} = \theta_{\max}$),
- a single beam per node,
- multiple beams per session,
- constraint on number of beams per node (possibly $> T$), and
- directional receiving antennas.

However, these are not addressed in this paper.

3.2 Energy Expenditure

In addition to RF propagation, energy is also expended for transmission (encoding, modulation, etc.) and reception (demodulation, decoding, etc.). We define:

- $p^T = \text{transmission processing power}$ and
- $p^R = \text{reception processing power.}$

We assume that these quantities are the same at all nodes, and we neglect any energy consumption occurring when the node is simply "on" without transmitting or receiving, although it would be straightforward to include it. However, for networks with regularly steady traffic, it is rare that a node will not be receiving or transmitting at any given time, especially in view of our assumption that multiple transceivers are available at each node. The total power expenditure of a node, when transmitting to a maximum range r over a sector of width θ, is

$$p = p^{RF}(r,\theta) + p^T + p^R \, 1(\text{node is not the source}), \quad (3)$$

where $p^{RF}(r,\theta)$ is defined in (2), and the indicator function is included because the p^R term is not needed for the source node. A leaf node, since it does not transmit but only receives, has a total power expenditure of p^R.

We assume that each node starts with a finite quantity of battery energy.[4] For example, Node i has energy $E_i(0)$ at time 0. The *residual energy* at Node i at time t is

$$E_i(t) = E_i(0) - \int_0^t P_i(\tau)d\tau, \quad (4)$$

where $P_i(\tau)$ is the total power expended at Node i at time τ.[5] We say that a node is "alive" as long as its residual energy is positive and that it dies when its residual energy decreases to zero. Based on our assumptions, a "dead" node cannot participate, even as a receive-only leaf node.

4 The Multicasting Problem

The establishment of a multicast tree requires the specification of the transmitted power levels, the frequencies used by each node, and the commitment of the needed transceiver resources throughout the duration of the session.

We assume that multicast session requests arrive to each of the N nodes at rate λ/N arrivals per unit time. The set of desired destinations is chosen randomly for each arrival. We say that a destination can be *reached* if the following conditions are satisfied:

- there exists a path from the source to it (i.e., the transmitted power required to support the path does not exceed $p_{\max}$ at any node),
- a transceiver is available (i.e., not already supporting another session) at each node along the path, and
- a suitable frequency assignment can be found to support the path (i.e., a noninterfering frequency is available to support the link between each node pair in the network along the path; these frequency assignments must not interfere with, or suffer interference from, currently ongoing sessions).

As noted earlier, all multicast requests are accepted as long as one or more of the intended destinations can be reached and paths are established to all reachable destinations, regardless of the cost required to do so.

4.1 Performance Measures

In this paper, we focus primarily on one particular performance measure, which is especially well suited for energy-limited applications, namely, the total delivered traffic volume during the lifetime of the network. We also consider the related quantity of traffic volume per unit energy.

We first introduce some notation. We assume that, once a session (multicast tree) is established, communication takes place at a constant rate of R bits/s, which is the same for each session request, and which is independent of λ. Session duration is exponentially distributed with mean $1/\mu = 1$.

Since partial multicast sessions may take place (because some nodes may be unreachable), the performance metric should provide a reward that reflects the number of destinations that are actually reached. We define

- $n_i = \#$ of intended destinations in session i,
- $m_i = \#$ of destinations reached in session i,
- P_i = sum of the transmitter powers used by all nodes in session i, and
- d_i = duration of session i.

4.1.1 Delivered Traffic Volume

The delivered traffic volume is directly proportional to both the number of destinations that are reached and to the duration of each session. Specifically, each destination node participating in multicast session i receives $b_i = R\, d_i$ bits during the course of the session. The total quantity of data delivered during session i is then

$$\begin{aligned} B_i &= \text{total number of bits delivered to all} \\ &\quad\ \text{reached destinations in session } i \\ &= m_i b_i. \end{aligned}$$

4. We assume that the battery has a fixed capacity, i.e., we neglect the fact that the total energy that can be supplied by a battery depends in part on the discharge rate and duty cycle [15]. We also neglect any nonlinear behavior, which may characterize power amplifiers especially at high output levels.

5. Since Node i may be transmitting as a member of several trees simultaneously, $P_i(\tau)$ is the sum of the powers for all such trees at time τ.

Then, the total quantity of information delivered to all destinations over an observation interval of X multicast requests is:

$$B_X^{total} = \sum_{i=1}^{X} B_i = R \sum_{i=1}^{X} m_i d_i. \tag{5}$$

4.1.2 Delivered Traffic Volume per Unit Energy

The energy expenditure in session i is $P_i d_i$. Thus, the total energy expenditure over the observation interval is

$$E_X = \sum_{i=1}^{X} P_i d_i. \tag{6}$$

Therefore, the delivered traffic volume per unit energy over an interval of X arrivals is

$$B_{X,E} = \frac{B_X^{total}}{E_X} = \frac{R \sum_{i=1}^{X} m_i d_i}{\sum_{i=1}^{X} P_i d_i}.$$

4.1.3 Local Cost Metrics

Tree formation consists of the specification of transmitting nodes and their downstream neighbors. When omnidirectional antennas are used, it is sufficient to specify the set of transmitting nodes and their RF transmission power levels; when directional antennas are used, the antenna pattern must also be specified. It is not feasible to find the multicast trees that guarantee the optimal values of global performance measures such as B^{total}, etc. Therefore, we have focused on the development of "local" strategies that depend on "local"[6] metrics, which find the multicast tree that attempts to minimize an appropriate cost function for each new multicast request.

In particular, the basic approach taken in [1] and [2] is to minimize the power needed to maintain the tree associated with each newly arriving session.[7] This power includes the RF transmission power of all transmitting nodes as well as the signal processing power expended at transmitting and receiving nodes. We recognize that local optimization does not guarantee global optimization, e.g., minimizing tree power does not guarantee the minimization of energy over an observation interval of many arrivals. Moreover, even if it were possible to do so, this would certainly not guarantee the optimization of the desired global performance measures. Nevertheless, it has been our experience that this approach works reasonably well.

The multicasting problem is similar to the broadcasting problem, except that only a specific subset of the nodes are required to be in the tree. It is well-known that the determination of a minimum-cost multicast tree in wired networks is a difficult problem, which can be modeled as the NP-complete Steiner tree problem, even though the broadcasting problem is easily formulated as the MST problem, which has low complexity. In wireless networks, even the broadcasting problem is NP-complete [7], [4]. Thus, heuristics are needed for both broadcasting and multicasting. The two basic approaches we have used for multicasting are the "pruning" of broadcast trees and the superposition of unicast paths [1], [2].

5 Algorithms for Broadcasting and Multicasting with Directional Antennas

We have considered two basic approaches for broadcasting and multicasting with directional antennas:

- Construct the tree by using an algorithm designed for omnidirectional antennas; then, reduce each antenna beam to the minimum possible width that can support the tree.
- Incorporate directional antenna properties into the tree-construction process.

The first approach can be used with any tree-construction algorithm. The "beam-reduction" phase is performed after the tree is constructed by using an additional "postprocessing" algorithm, which is appended to the tree-construction algorithm. The second approach, which requires decisions on beamwidth to be made at each step of the tree construction process, can be used only with algorithms that construct trees by adding one node at a time, such as BIP (and its multicasting counterpart MIP). In this section, we describe these approaches in detail.

5.1 An Approach Based on Beamwidth-Reduction: Reduced Beam BIP (RB-BIP) and MIP (RB-MIP)

First, a low-cost broadcast or multicast tree is formed, using any tree-construction algorithm (e.g., BIP or MIP), under the assumption that the transmitting antennas are omnidirectional. Then, after the tree is constructed in this manner, each transmitting node's antenna beamwidth is reduced to the smallest possible value that provides coverage of the node's downstream neighbors, subject to the constraint $\theta_{\min} \leq \theta \leq 360$. Thus, the tree structure is independent of $\theta_{\min}$. We assume perfect antenna patterns that provide uniform gain throughout the cone of beamwidth θ (with no sidelobes), so it is not necessary to extend θ beyond the direction of the nodes at the edges of the cone. When applied to BIP, the resulting scheme is called Reduced-Beam BIP (RB-BIP); when applied to MIP, the resulting scheme is called RB-MIP.

As noted in Section 3.1, we assume that each node knows the location of all of its potential neighbors (at $p^{RF} \leq p_{\max}$), and can thus point its beam in the appropriate direction to provide connectivity to all of its children in the tree. Fig. 2a shows Node s with its children (as determined by BIP). Once the set of children is determined, the communication range r is the distance to the most-distant child. To minimize the required power (for a given range), the minimum beamwidth that reaches all children must be determined.[8] This problem involves finding the largest sector that can be omitted from omnidirectional coverage, while still reaching all of the children. In Fig. 2a, the choice of the appropriate sector of coverage (which excludes the sector of width $360 - \theta$) is obvious from the graphical

6. "Global" is used here to refer to optimization over a long observation interval. "Local" is used here both in the sense of time-local (i.e., for each arrival of a multicast session request), as well as in the topological sense (i.e., pertaining to an individual link or node).

7. In Section 6, we introduce a cost metric that also involves the residual energy at each node.

8. In some applications, it may be desirable to incorporate interference considerations into the choice of the sector that is to be omitted; however, we do not do so in this paper.

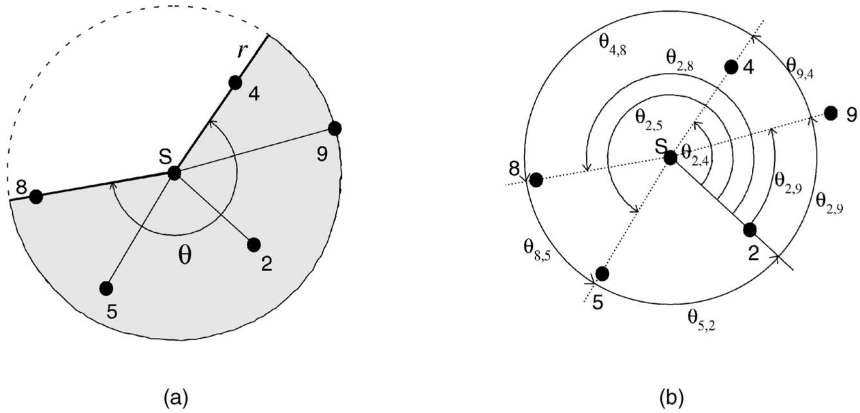


Fig. 2. Choosing the sector to be covered by the beam. (a) Sector of beamwidth θ and (b) angles used to determine sector.

representation. However, it is helpful to discuss the algorithmic procedure used to determine this sector.

Given the coordinates of the nodes, we would like to determine the sector of minimum beamwidth that includes all children nodes. In general, Node s has n children, which are a subset of the nodes in $\{1, 2, \dots, N\}$. In the example of Fig. 2, $n = 5$ and the children are Nodes {2, 4, 5, 8, 9}. Using the ray from Node s to its neighbor with lowest ID (Node 2) to define the reference direction,[9] we measure the angles $\{\theta_{2,4}, \theta_{2,5}, \theta_{2,8}, \text{ and } \theta_{2,9}\}$ (all in the same, e.g., counter-clockwise, direction), as shown in Fig. 2b. The complexity of this operation is $O(n)$ since Node s has n children. We then sort these angles in increasing order, which in this example is: $\{\theta_{2,9}, \theta_{2,4}, \theta_{2,8}, \theta_{2,5}\}$, thereby determining that the "cyclic" ordering of the nodes is: {2, 9, 4, 8, 5, 2}. This sorting process can be done with complexity $O(n \log n)$ using heap sort [16]. Once this is done, we can determine the nonoverlapping angles that define the sectors between adjacent nodes, namely, $\{\theta_{2,9}, \theta_{9,4}, \theta_{4,8}, \theta_{8,5}, \text{ and } \theta_{5,2}\}$. We then find the largest such angle, which in this case is $\theta_{4,8}$ (resulting in an angle of coverage of $\theta = 360 - \theta_{4,8}$). This operation has complexity $O(n)$. Therefore, the overall complexity of the sector-choosing algorithm for a node with n children (once the node's children have been selected) is

$$O(n) + O(n \log n) + O(n) = O(n \log n).$$

Finally, since the number of children a node may have is at most $N - 1$, the overall complexity is $O(N \log N)$ at each transmitting node. Since there are $O(N)$ transmitting nodes, the beam reduction operation (which is done once at each transmitting node after the tree has been constructed, e.g., by using BIP) has overall complexity $O(N^2 \log N)$. Since BIP has complexity $O(N^3)$ (see Section 5.3), adding the beam-reduction operation (as a postalgorithm process) does not affect its overall degree of complexity, and the complexity of RB-BIP remains $O(N^3)$.

5.2 An Approach Based on Incremental Power: Directional BIP (D-BIP) and MIP (D-MIP)

In [1] and [2], we proposed the Broadcast Incremental Power (BIP) algorithm, a centralized heuristic for the determination of low-power broadcast trees in networks with omnidirectional antennas. BIP is the basis for the Multicast Incremental Power (MIP) algorithm, under which the tree produced by BIP is pruned by eliminating all transmissions that are not needed to reach the members of the multicast group. More specifically, under MIP, nodes with no downstream destinations do not transmit, and some nodes may be able to reduce their transmitted power (i.e., if their more-distant downstream neighbors have been pruned from the tree).

BIP is similar in principle to Prim's algorithm for the formation of minimum-cost spanning trees (MSTs), in the sense that new nodes are added to the tree one at a time (on a minimum-cost basis) until all nodes are included in the tree. In fact, the implementation of this algorithm is based on the standard Prim algorithm, with one fundamental difference. Whereas the inputs to Prim's algorithm are the link costs p_{ij} (which remain unchanged throughout the execution of the algorithm), BIP must dynamically update the costs at each step (i.e., whenever a new node is added to the tree). This updating is done to reflect the fact that the cost of adding a new node to a transmitting node's list of neighbors is the *incremental cost*, i.e., the additional cost associated with adding a new downstream neighbor, given that the node is already transmitting at some particular power level. Consider an example in which Node i is already in the tree (it may be either a transmitting node or a leaf node) and Node j is not yet in the tree. If Node j is already participating in T sessions (hence, no transceivers are available for an additional session), the cost of adding it to the tree is set to ∞.[10] Otherwise, for all such Nodes i (i.e., all nodes already in the tree) and Nodes j (i.e., nodes not yet in the tree), the following is evaluated:

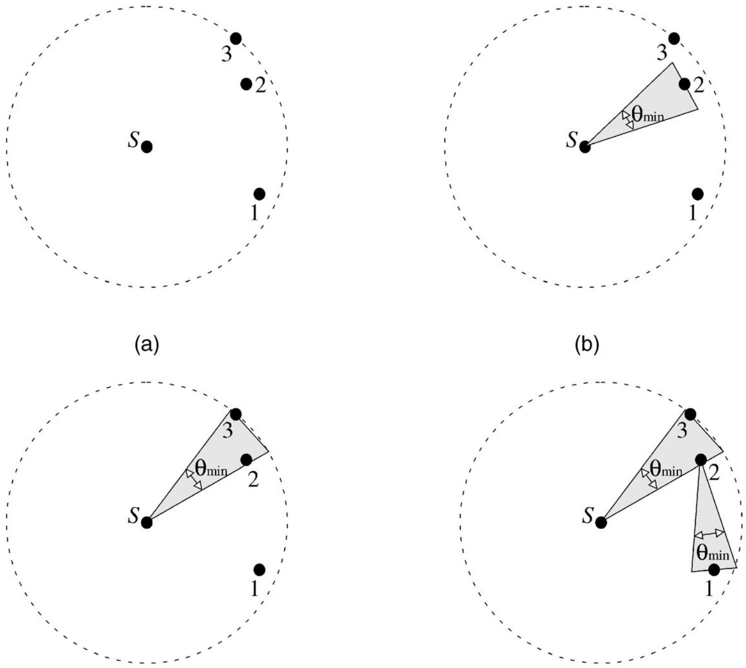
$$p'_{ij} = p_{ij} - p_i, \tag{8}$$

where p_{ij} is the link-based cost (power) of a transmission[11] between Node i and Node j (i.e., it is $r_{ij}^{\alpha} + p^T$), and p_i is Node i's transmission cost prior to the addition of Node j; (which includes p^T if node i is already transmitting; if Node i is currently a leaf node, $p_i = 0$). The quantity p'_{ij} represents the *incremental cost* associated with adding Node j to the set of nodes to which Node i already transmits. The pair $\{i, j\}$ that results in the minimum value of p'_{ij} is selected, i.e., Node i transmits at a power level sufficient to reach Node j.

9. Any child node could have been used to determine the reference direction.

10. It is also possible to associate a higher cost with nodes that have low "residual capacity" (i.e., few available transceivers); however, we do not do so in this paper.

11. We neglect p^R in this cost measure because it is the same for all possible Node js. However, p^R is included when energy consumption is evaluated.



(a) (b) (c) (d)

Fig. 3. Addition of nodes in D-BIP. (a) Source node and three destination nodes. (b) Add Node 2 to tree. (c) Shift beam and increase power to add Node 3. (d) Use Node 2 as relay to reach Node 1.

Thus, one node is added to the tree at every step of the algorithm.

The incremental power philosophy, originally developed for use with omnidirectional antennas, can be applied to tree construction in networks with directional antennas as well. At each step of the tree-construction process, a single node is added, as above. However, whereas the only variable involved in computing the cost (and incremental cost) in the omnidirectional case was the transmitter power, the directional-antenna case involves the choice of beamwidth θ as well. Based on the propagation model of (2), the required RF power increases in proportion to the α power of the distance to the farthest downstream neighbor and linearly with θ.

Consider a situation in which Node i is already transmitting to several other nodes. The incremental cost of adding Node j to Node i's set of downlink neighbors depends on the relative location of Node j with respect to the region already included in Node i's antenna's cone of coverage. For example, if Node j is located within the angle of Node i's beam, it suffices to increase Node i's communication range, without changing the width or direction of the beam.[12] On the other hand, if Node j is not located within the angle of Node i's beam, then the beam must be adjusted; this is usually done by increasing θ, although it is sometimes possible to simply shift the beam if all of a node's downstream neighbors are located within a cone not greater than $\theta_{\min}$. Thus, to add a new node, it is sometimes sufficient to simply increase transmission range, it is sometimes sufficient to simply shift the beam, sometimes the beam has to be made wider, and sometimes a combination of increased communication range and beam characterization must be done. Note that there is no incremental cost associated with shifting a beam (while maintaining the same angle of coverage).

Fig. 3a shows a simple example in which the source node S wants to add nodes 1, 2, and 3 to the tree. Node 2 is the closest to S, so it is added first (as in standard omnidirectional BIP); in Fig. 3b, we show a beam of width $\theta_{\min}$ that is centered on the ray between S and Node 2. We must then decide which node to add next (Node 1 or Node 3), and which node (that is already in the tree) should be its parent. In this example, the distance between S and Node 3 is only slightly greater than that between S and Node 2; additionally, the beam from S to Node 2 can be shifted to include both Node 2 and Node 3, without increasing its beamwidth. Therefore, Node 3 is added next by increasing the communication range of S and by shifting the beam (see Fig. 3c). Finally, Node 1 must be added to the tree. One possibility is for S to simply increase its beamwidth, resulting in a situation similar to that in Fig. 2 (it would not have to increase its range because its distance to Node 1 is less than its distance to Node 3). However, doing so would result in a huge increase in beamwidth (we noted above that RF power is directly proportional to beamwidth). In our example, it is more cost-effective for Node 2 to become the parent of Node 1 (using a beamwidth of $\theta_{\min}$, as shown in Fig. 3d). Note that the tree produced by this algorithm depends on both $\theta_{\min}$ and the propagation

12. It is also necessary to examine whether Node j could be added to the tree at lower cost by using a different node (e.g., one of Node i's downstream neighbors) as its upstream neighbor.

constant α. By contrast, the tree produced by RB-BIP is independent of $\theta_{\min}$, although it does depend on α.

When applied to the broadcasting problem, the resulting scheme is called Directional BIP (D-BIP). When applied to the multicasting problem, the first step is to form a broadcast tree using D-BIP. To implement Directional MIP (D-MIP), the broadcast tree produced by D-BIP is pruned, as discussed at the beginning of this section. Note that when $\theta_{\min} = 360$, D-BIP, RB-BIP, and BIP are identical.

In [1], [2], we described the "sweep" operation, in which redundant transmissions, as well as transmissions that can be reduced in power without compromising the tree structure, are discovered. Use of the sweep was shown to provide modest improvement in energy efficiency in networks that use omnidirectional antennas. The sweep operation can be performed either after the tree is constructed (an approach suitable for all tree-construction algorithms) or at each step of the tree-construction algorithm (an approach feasible only for those algorithms, such as BIP, that add a single node to the tree at every step). Typically, more improvement is observed when the former approach is used (in omnidirectional applications) because the impact of each change resulting from the sweep affects a greater number of nodes.

However, a fundamental characteristic of the sweep makes it unattractive in applications with directional antennas. In omnidirectional applications (and, hence, when the RB versions of tree-construction algorithms are used as well), the sweep often results in an increase in the number of one or more nodes' downstream neighbors. Consequently, there is often an increase in the required angle of coverage needed by the antenna in RB algorithms and, thus, in an increase in that node's RF power, and possibly in the overall tree power. With D-BIP and D-MIP, this problem does not arise; however, there are fewer situations in which the sweep will discover improvements (because widening the beam will often result in increased power). Therefore, the performance results presented in this paper do not make use of the sweep.

5.3 Complexity of BIP and D-BIP

First, let us consider the complexity of BIP with omnidirectional antennas. One node is added to the tree at each step; thus, N steps are needed to construct a tree with N nodes. Consider step k, at which k nodes are already in the tree (we call them "inside" nodes) and the remaining $(N-k)$ "outside" nodes must eventually be added. Thus, the number of inside nodes is $O(N)$, and the number of outside nodes is also $O(N)$. At each step, the algorithm must determine which node can be added at least incremental cost. Thus, it is necessary to know the cost from each inside node to every outside node, i.e., for each inside node, we must find the least-cost outside node; we then take the minimum over all inside nodes.

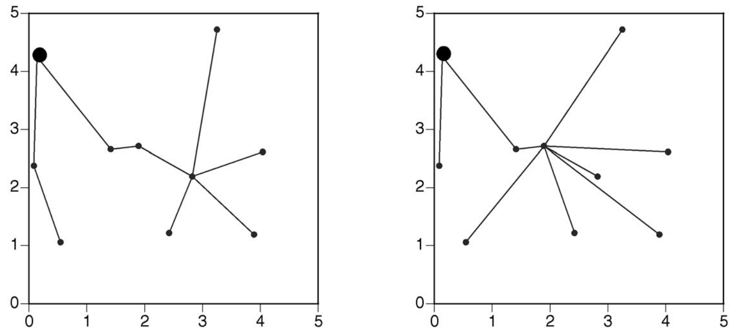
Most of the database needed to implement BIP remains unchanged from one step to the next. The only costs that must be updated are those that relate to the new child node (which changes its status from an outside node to an inside node) and to its parent. The cost between the new child node and all of the remaining outside nodes must be computed; the complexity of this operation is $O(N)$. Additionally, since its parent has increased its RF power, the cost from the parent to all remaining outside nodes must be updated; this also has complexity $O(N)$. For the other inside nodes, we simply delete the entry corresponding to the node that was just added to the tree; this also has complexity $O(N)$. Therefore, updating the database requires three sequential operations of complexity $O(N)$, which results in a complexity of $O(N)$. Once the database is updated, the minimum-cost new node must be chosen. The complexity of this search is at most $O(N^2)$. Therefore, the complexity at each step is at most $O(N^2)$. Since there are N steps, the overall complexity is at most $O(N^3)$.

Now, let us consider the complexity of D-BIP, for which we must incorporate the impact of the choice of antenna beam angle at *each* step (in contrast to the case of RB-BIP, in which the beam reduction process was executed once at each transmitting node, after the tree was constructed). As in the case of omnidirectional BIP, most of the database remains unchanged from one step to the next. The recomputation of costs to outside nodes is necessary only for two nodes, namely, the parent of the node that was just added and for the child (since it becomes an inside node). First, consider the parent. Following the discussion of Section 5.1, the complexity of choosing the best outside node for any given inside node (which involves choosing the antenna beamwidth θ as well) is $O(N^2 \log N)$. Now, consider the child. It is necessary to compute its cost to all outside nodes. Since the child doesn't have any children of its own yet, it is not necessary to compute the beamwidth θ because it is $\theta_{\min}$ whenever a node has a single child. Thus, the complexity of computing the cost from the child to all remaining outside nodes $O(N)$. For every other inside node, the newly added node is simply deleted from its list of outside nodes, without recomputing anything. Therefore, the complexity of each step is $O(N^2 \log N)$ and since there are N steps, the overall complexity of D-BIP is $O(N^3 \log N)$.

5.4 Example Broadcast Trees

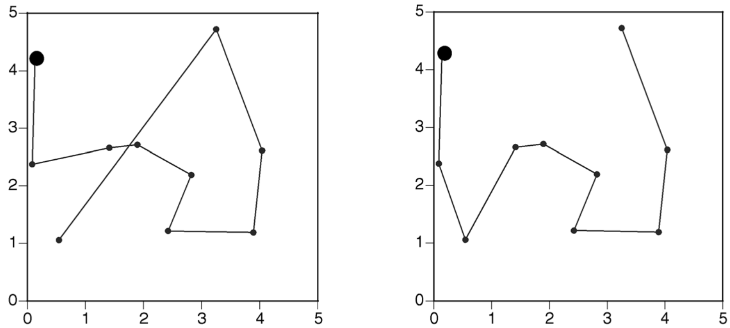
Fig. 4a shows the broadcast tree produced by BIP for a 10-node network, where the source node is shown larger than the other nodes. As noted in Section 3.1, RB-BIP uses the same tree as BIP (which is based on omnidirectional antennas); the only difference is that the antenna beamwidth is reduced. Fig. 4b shows the optimal tree for omnidirectional antennas, which was obtained by exhaustive search. The tree structure, as well as the resulting value of total tree power P, depend on the value of the propagation constant α; our results are based on $\alpha = 2$. Tree power P is listed in the figure caption for $\theta_{\min} = 360$ (the omnidirectional case), as well as $\theta_{\min} = 30$ and 1. There is relatively little power savings when $\theta_{\min}$ is reduced below 30 because the two highest-range transmissions require the use of $\theta > 30$ to reach all of their downstream neighbors.

Under D-BIP (unlike RB-BIP), the tree structure depends on the value of $\theta_{\min}$. Fig. 5a shows the tree for the same network for D-BIP with $\theta_{\min} = 1$. In this example, D-BIP produces a tree in which each node has only a single downstream neighbor (thus, $\theta = \theta_{min}$ at each node) resulting in a zigzag path with no branching. The value of P is greatly reduced by using highly directional antennas. However, this value is 84 percent greater than that of the optimal tree for $\theta_{\min} = 1$, as shown in Fig. 5b.



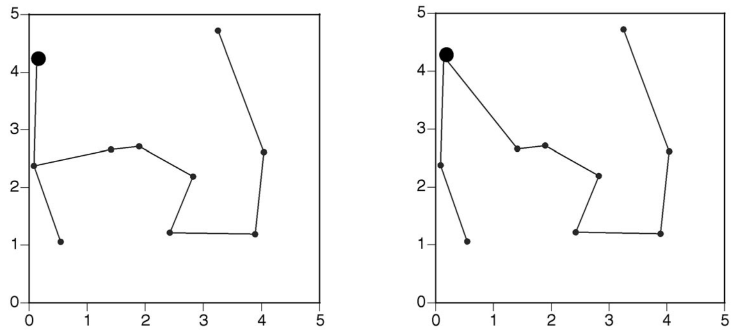
(a) (b)

Fig. 4. Example 10-node broadcast trees based on use of omnidirectional antennas (the same tree is used for RB-BIP, independent of the value of $\theta_{\min}$). (a) BIP, RB-BIP: $\theta_{\min} = 360 : P = 14.06$, $\theta_{\min} = 30 : P = 4.26$, and $\theta_{\min} = 1 : P = 3.99$. (b) optimal (based on $\theta_{\min} = 360$). $\theta_{\min} = 360 : P = 10.71$, $\theta_{\min} = 30 : P = 3.728$, and $\theta_{\min} = 1 : P = 3.709$.



(a) (b)

Fig. 5. Example 10-node broadcast trees for $\theta_{\min} = 1$. (a) D-BIP ($P = 0.1051$) and (b) optimal ($P = 0.05707$).



(a) (b)

Fig. 6. Example 10-node broadcast trees for $\theta_{\min} = 30$. (a) D-BIP ($P = 1.722$) and (b) optimal ($P = 1.607$).

Fig. 6a shows the D-BIP tree for the same network, but with $\theta_{\min} = 30$. Here, P is 7.2 percent greater than that of the optimal tree, as shown in Fig. 6b.

These results demonstrate that the use of directional antennas can facilitate considerable energy saving through the use of algorithms such as RB-BIP and D-BIP. Moreover, D-BIP provides lower-power trees than RB-BIP for a given value of $\theta_{\min}$, and this advantage increases as $\theta_{\min}$ decreases. However, when $\theta_{\min}$ is very small, even the tree produced by D-BIP is likely to have a significantly higher value of RF transmission power than the optimal tree (on a percentage basis).

We attribute the relatively good performance of BIP when $\theta_{\min} \geq \theta$ (as measured by the closeness of tree power to its optimal value, on a percentage basis) to the wireless multicast advantage (see Section 3.1). However, this property no longer applies when highly directional antennas are used because power is directly proportional to beamwidth θ; thus, it is costly to expand a beam to accommodate additional nodes. Therefore, the greedy nature of our incremental power approach suffers when used with extremely narrow beams, and alternative approaches may be desirable.

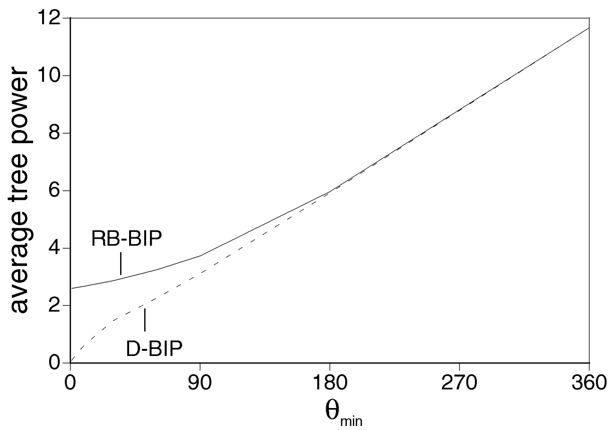


Fig. 7. Average RF tree power for broadcasting examples with a 50-node network using BIP, based on use of directional antennas ($p_{max} = 25, \alpha = 2$).

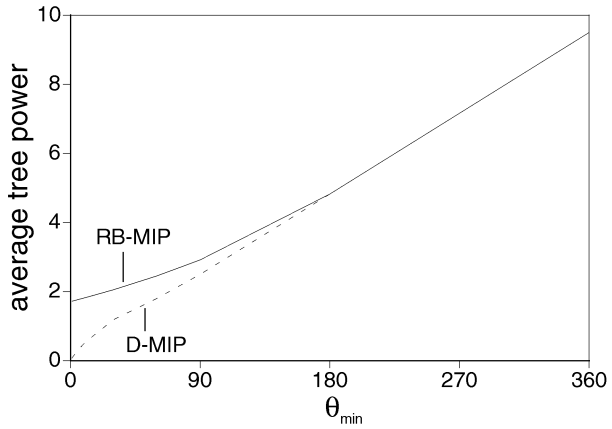


Fig. 8. Average RF tree power for multicasting examples with a 50-node network using MIP, based on use of directional antennas ($p_{max} = 25, \alpha = 2$).

5.5 Average Tree Power

We have simulated BIP and MIP using the two directional antenna schemes; thus, we have simulated a total of four algorithms, namely, RB-BIP, D-BIP, RB-MIP, and D-MIP. In all cases, we consider a network of $N = 50$ nodes that are randomly located in a region with dimensions 5×5 (arbitrary units of distance); the same node locations are used in our broadcasting and multicasting examples. We let $\alpha = 2$, i.e., the required RF power value is r^2. RF transmission power levels are bounded by $p_{min} = 0$ and $p_{max} = 25$ (corresponding to a maximum communication range of 5). The results presented in this section are based on the availability of unlimited numbers of transceivers, frequencies, and energy; thus, it is always possible to include all destinations in the tree. We are thus able to focus on the impact of adaptive beamwidth capabilities on energy efficiency. We also neglect the impact of signal-processing power by setting $(p^T, p^R) = (0, 0)$. Our results are based on simulation runs, each consisting of $X = 10,000$ multicast sessions.

First, we consider broadcasting. The source node is chosen at random for each session, and RB-BIP or D-BIP is used to construct the broadcast tree based on that source node. Fig. 7 shows the average tree RF power as a function of $\theta_{\min}$ for RB-BIP and D-BIP. As expected, both schemes provide substantial reduction in RF power as compared to omnidirectional antennas when $\theta_{\min} \ll 360$ because energy is concentrated where it is needed. For $\theta_{\min} \geq 180$ there is vitually no difference in performance between the two schemas. However, for small values of $\theta_{\min}$, D-BIP provides a substantial reduction in RF power as compared with RB-BIP.

Fig. 8 shows average tree power for multicasting examples in which RB-MIP and D-MIP are used. As above, the source node is chosen at random for each session. Multicast groups are chosen randomly for each session request; the number of destinations is uniformly distributed between 1 and $N - 1$. Results are qualitatively similar to those for the broadcasting case, although the values of tree power are somewhat lower. The power of a multicast tree generated by MIP is always less than or equal to that of a broadcast tree generated by BIP because MIP forms multicast trees by pruning broadcast trees generated by BIP.

In view of the similarity (both qualitative and quantitative) of the results for broadcasting and multicasting, we focus on multicasting in the rest of this paper.

6 The Incorporation of Resource Limitations

The discussions in the previous sections implicitly assume that sufficient resources are available to implement the trees created by the algorithms. These resources include transceivers, frequencies, and battery energy. In this section, we discuss how limitations on these resources are incorporated into our model and how our algorithms can be modified to cope with limited energy.

It is straightforward to incorporate the impact of a finite number of transceivers. When constructing a tree for a new arrival, the cost of a node is set to ∞ if all of its transceivers are currently supporting other sessions. However, the modeling of finite frequency resources is much more complicated.

6.1 Bandwidth Limitations

Let us consider the case in which Node *m* wants to transmit to Node *n*. Any particular frequency *f* may be unusable for one of the following reasons:

- *f* is already in use (for either transmission or reception) at either Node *m* or Node *n*,
- *f* is being used by one or more nodes that create interference at Node *n*, thereby preventing reception at *f*, and
- the use of *f* by Node *m* would interfere with ongoing communications at other nodes.

In this paper, we use the following basic greedy approach for frequency assignment, which we referred to as FA1 in [13]:

Assume the availability of an infinite number of frequencies when forming the tree (the approach used in [1] and [14]). Then attempt to assign the available frequencies to the tree. The assignment process is complete when either frequencies have been assigned to all transmis-

sions or when no additional frequencies are available to support portions of the tree.

Under this scheme, the tree construction process ignores the possibility that frequencies may not be available to provide the required connectivity. Thus, if appropriate frequencies cannot be found along the paths to all desired destinations, then some destinations will not be reached. We have used a greedy version, in which frequencies are assigned using an orderly procedure, without the possibility of backtracking to change assignments and without the use of exhaustive search (or other scheme) to determine whether a consistent frequency assignment is possible. Specifically, we simply assign the lowest-numbered available noninterfering frequency to each node. Thus, this scheme can result in unreached destinations, even though they might be reachable through a better frequency assignment. But, this is a common characteristic of all heuristic procedures.

In [13], we also considered an alternative scheme (FA2) under which, at each step of the tree-construction, the frequency is chosen along with the transmission power level. Under FA2, the tree is formed using only nodes that do, in fact, have frequencies available. Again, there is no guarantee that all destinations will be reached. However, FA2 provides a richer search space than FA1. In this paper, we focus exclusively on FA1 because it is simple to use and is applicable to any tree-construction algorithm. FA2 can be used with BIP (and similar schemes in which one node is added to the tree at each step), but not with some of the other algorithms discussed in [1] and [2].

6.2 Energy Limitations

Use of a cost metric that involves only the total power required to maintain the tree can result in rapid energy depletion at some nodes. When nodes "die" in this manner, it may be no longer possible to create energy-efficient trees.

We can discourage the inclusion of energy-poor nodes in the multicast tree by increasing the cost associated with their use. In (4), we defined the residual energy at Node i at time t to be $E_i(t)$. We now define the cost of a link between Node i and Node j to be

$$C_{ij} = p_{ij}\left(\frac{E_i(0)}{E_i(t)}\right)^{\beta}, \tag{9}$$

where β is a parameter that reflects the importance we assign to the impact of residual energy.[13] Clearly, when $\beta = 0$, the link cost is simply the power needed to maintain the link.

The incremental cost associated with adding Node j to the set of Node i's downstream neighbors, given that Node i is already transmitting at power level p_i (hence, at cost C_i) is:

$$C'_{ij} = C_{ij} - C_i. \tag{10}$$

When β is too small, too much emphasis may be placed on the construction of energy efficient trees, resulting in the rapid depletion of energy at some of the nodes. By contrast, when β is too large, too much emphasis may be placed on balancing energy use throughout the network, while underemphasizing the need for energy efficiency.

Performance results in Section 7 and [8] show the beneficial effects of using β in the range [0.5, 2]. It would be possible to develop alternative cost functions to (9) that also discourage the use of energy-poor nodes; we make no claim of optimality. Our objective is to demonstrate that load balancing based on residual energy can extend a network's useful lifetime.

7 Performance Results

Important performance measures for energy-constrained networks include network lifetime and delivered traffic volume. In this section, we present our performance results for the two multicast tree-construction algorithms we have developed for directional antennas, namely, Reduced-Beamwidth MIP (RB-MIP) and Directional MIP (D-MIP).

As in Section 5.5, we have simulated the performance of RB-MIP and D-MIP for a network of $N = 50$ nodes that are randomly located in a region with dimensions 5×5; the same node locations are used in all examples presented in this paper. In extensive performance evaluation, we have observed that these results are representative of other random node distributions as well. We present results primarily for a propagation constant value of $\alpha = 2$, which results in required RF power values of r^2 to support a link between two nodes that are separated by distance r; we also present a limited set of results for $\alpha = 4$. We set arbitrary values for transmission processing power (p^T) and reception processing power (p^R). In particular, we consider $(p^T, p^R) = (0, 0)$ as well as "moderate" (0.01, 0.1) and "high" (0.1, 1) values of these quantities. RF transmission power levels are bounded by $p_{\min} = 0$ and $p_{\max} = 25$. In most of our experiments, the initial energy at each node is 200 (arbitrary units, consistent with the units of distance).[14] We demonstrate the impact of incorporating residual energy into the cost metric and compare performance for $\beta = 0$, 0.5, 1, and 2.

In our simulations, multicast requests arrive with interarrival times that are exponentially distributed with rate λ/N at each node; we have used $\lambda = 1$ in our simulations. Session durations are exponentially distributed with mean 1. Multicast groups are chosen randomly for each session request; the number of destinations is uniformly distributed between 1 and $N - 1$.

Each simulation run consists of $X = 10,000$ multicast sessions, some of which may be blocked because of lack of resources (which in general include transceivers, frequencies, and energy). The same random number sequence is used to drive each of our experiments, thereby facilitating a meaningful comparison of results for different values of β.

7.1 Network Lifetime

A fundamental issue in limited-energy applications is network lifetime, i.e., the interval over which the network can

13. Residual energy was incorporated into the cost metric in a similar manner in [17].

14. We assume that if a node is alive at the beginning of a session, it will be able to complete the session (regardless of whether it is a transmitting or a receive/only node). Thus, we neglect the minor "end effects" associated with a node's death during a session.

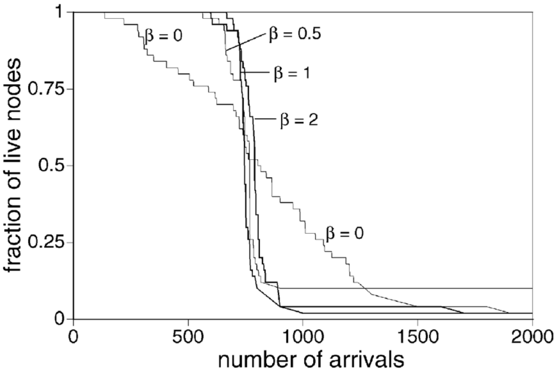


Fig. 9. Evolution of number of live nodes under MIP with omnidirectional antennas for 50-node network (zero processing power).

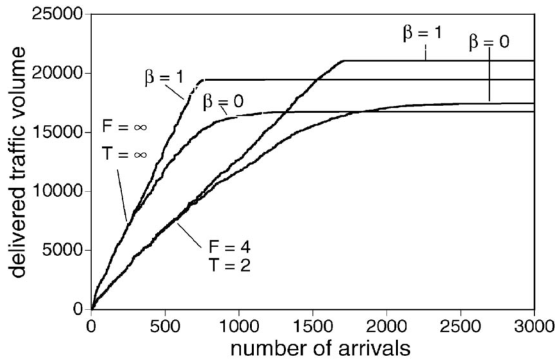


Fig. 10. Evolution of cumulative bit volume under MIP for several sets of (F, T) pairs $(\beta = 0; (p^T, p^R) = (0, 0))$.

provide acceptable levels of service. Clearly, a suitable definition of network lifetime depends on the specific application. For example, in some applications, one may view network death as the time at which the first node dies (e.g., see [17]) because it is no longer possible to reach all of the nodes. Alternatively, network death may be defined as the death of a specified fraction of the nodes. In this paper, we don't specify a particular definition of network death, although we do feel that a reasonable definition of acceptable performance would require that at least 50 percent of the nodes remain alive. Instead, we examine the time evolution of the number of live nodes.

In this section, we consider the case of unlimited numbers of transceivers and frequencies, but finite energy at each node. We present results for omnidirectional antennas (although results are qualitatively similar for directional antennas). Thus, we are able to focus on the impact of energy constraints, without addressing other system parameters. In such cases, all desired destinations can be reached, provided that live nodes are available to support the required trees.

Fig. 9 shows the evolution of the number of live nodes as a function of the number of session arrivals for $\beta = 0$, 0.5, 1, and 2. Results are shown for the case of zero processing power, i.e., $(p^T, p^R) = (0, 0)$. As noted in Section 6, the use of nonzero values of β tends to discourage the use of nodes that have little residual energy. The use of $0.5 \leq \beta \leq 2$, rather than 0, results in a significant delaying of the first node's death, and keeps a large fraction (e.g. 80 percent or 90 percent) of the nodes alive for a considerably greater number of sessions. Specifically, for zero precessing power, when $\beta = 0$, the first node dies at arrival 136; for $\beta = 0.5$, 1, and 2, the first node dies at arrival 563, 668, and 599, respectively. Results are qualitatively similar when $(p^T, p^R) = (0.1, 1)$, except that nodes die much faster because of the energy consumed by signal processing (see [8]).

Moreover, for $0.5 \leq \beta \leq 2$, once about 10 percent of the nodes have died, the fraction of live nodes decreases to below 10 percent shortly thereafter. The rapid death of nodes in this manner is not a harmful effect. Once about 50 percent of the nodes are dead, a significant number of the remaining live nodes are typically unreachable. Thus, the fact that the use of $\beta = 0$ maintains a certain fraction (say 25 percent) of the nodes alive considerably longer than use of larger values of β is not beneficial.

Thus, for $0.5 \leq \beta \leq 2$, we have achieved a high degree of load balancing that keeps almost all of the nodes alive for a relatively long time, thereby maintaining network connectivity and high levels of throughput much longer than for the case in which $\beta = 0$. In view of the relative insensitivity of node lifetime to the value of β (in the region $0.5 \leq \beta \leq 2$), we use $\beta = 1$ in the examples presented in this paper. No claim for optimality is made.

Since we use a finite value of $p_{\max}$, it is typical to achieve a final state in which a number of nodes still have energy, but further communication is impossible because of a lack of connectivity among the live nodes.[15]

7.2 Delivered Traffic Volume

We now consider the delivered traffic volume B^{total}. In doing so, we address the impact of realistic constraints on the number of transceivers (T) available at each node and on the number of frequencies (F) available for communication. Our modeling assumptions are the same as those of the previous section. Unlike the case of infinite transceiver and frequency resources, performance depends strongly on the arrival rate λ because high traffic loads require a large number of transceivers and frequencies to support them. We present results for MIP, first for omnidirectional and then for directional antennas. Our results are based on the use of frequency assignment scheme FA1. As noted in Section 5.2, the "sweep" is not used.

7.2.1 Omnidirectional Antennas

Fig. 10 shows the time evolution of B^{total} under MIP, with omnidirectional antennas. To illustrate the impact of a finite number of transceivers and frequencies, we present results for $(F, T) = (4, 2)$ and (∞, ∞). To illustrate the impact of using a nonzero value of β, we present results for $\beta = 0$ and 1. These results are based on an arrival rate of $\lambda = 1$ and $(p^T, p^R) = (0, 0)$. One unit on the vertical axis corresponds to the delivery of a message of average length (one time unit) to a single destination (see definition in (5)). The initial value of energy at each node is $E_i(0) = 200$.

Consider first the curves for $\beta = 0$. The use of $(F, T) = (4, 2)$ results in a significantly slower rise in the value of delivered traffic volume; the blocking of destinations (or

15. If we had included the energy consumption associated with operation when a node is simply "on," but neither transmitting nor receiving, all nodes would have died eventually.

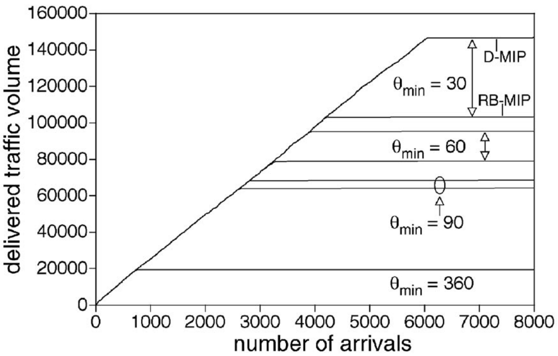


Fig. 11. Evolution of cumulative bit volume under MIP with directional antennas for D-MIP and RB-MIP; $T = \infty$, $F = \infty$ ($\beta = 1; (p^T, p^R) = (0, 0)$).

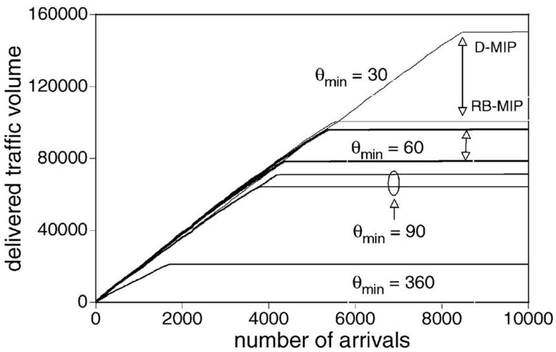


Fig. 12. Evolution of cumulative bit volume under MIP with directional antennas for D-MIP and RB-MIP; $T = 2$, $F = 4$ ($\beta = 1; (p^T, p^R) = (0, 0)$).

perhaps even entire sessions) results from the unavailability of transceiver or frequency resources at one or more nodes. Nevertheless, the final value (i.e., the total delivered traffic volume) is 4.16 percent greater than that for the case of $(F, T) = (\infty, \infty)$. Overall, the value of B^{total} is relatively insensitive to the values of *F* and *T*, even though these parameters have a significant impact on the *rate* of traffic delivery. A possible explanation for the higher value of B^{total} that is achieved for low values of *F* and *T* is that costly destinations are more likely to be blocked when few resources are available, thus resulting in a lowered average cost per destination that is actually reached.

Now, let us consider the impact of setting $\beta = 1$. For any (*F,T*) pair, the curve can be approximated well by a linear increase until the final value is reached, a departure from the asymptotic performance observed for $\beta = 0$. This behavior can be explained by the fact that the load balancing that results from the use of $\beta = 1$ provides a rapid transition from a state in which most nodes are alive to one in which most are dead, as shown in Fig. 9. Thus, there are two distinct regions of operation. When all (or most) nodes are alive, the rate of traffic delivery is maintained at (or near) its maximum value. When most nodes are dead, the rate of traffic delivery is close to (or equal to) zero. We again observe that, although $(F, T) = (4, 2)$ provides a lower initial rate of traffic delivery, it provides an 8.22 percent higher value of B^{total}. Additional discussion of delivered traffic volume in energy-constrained systems with omnidirectional antennas may be found in [8].

7.2.2 Directional Antennas

We now consider the case of directional antennas. Fig. 11 shows the time evolution of B^{total} for RB-MIP and D-MIP for several values of $\theta_{\min}$. Results are shown for $\beta = 1$, zero processing power, and $T = F = \infty$. The case of $\theta_{\min} = 360$ corresponds to the use of omnidirectional antennas. Our first observation is that the use of RB-MIP and D-MIP provide significantly increased values of delivered traffic volume, and that this volume increases as $\theta_{\min}$ decreases. The increase is less than linear in $1/\theta_{\min}$ because some beamwidths may be greater than $\theta_{\min}$.

For $\theta_{\min} = 30$, 60, and 90, two curves are shown for each value; the lower curve is for RB-MIP and the upper curve is for D-MIP. In all cases, D-MIP provides better performance than RB-MIP, and its advantage increases as $\theta_{\min}$ decreases. As in the omnidirectional case for $\beta = 1$, the curves can be closely approximated by straight lines with slope independent of $\theta_{\min}$, until the final value is reached. For the present case of unlimited transceiver and frequency resources, this observation is not surprising. Since the network has sufficient resources to deliver all offered traffic, and since the use of $\beta = 1$ achieves good load balancing (causing the nodes to die at almost the same time), the delivered traffic volume increases essentially linearly until the nodes die. Reducing the value of $\theta_{\min}$ results in a reducing of the energy expenditure, while not affecting the rate of traffic delivery (all offered traffic while the nodes are alive, and none while they are dead). However, the point at which the nodes die (as a function of $\theta_{\min}$) cannot be predicted without actually running the simulation.

Fig. 12 shows similar results for finite transceiver and frequency resources, namely $(F, T) = (4, 2)$. As in the case of infinite resources, each curve shows a linear increase until its final value is reached; however, there are noticeable differences in slope for different values of $\theta_{\min}$. As a consequence of the reduced resources, a significant fraction of destinations is blocked, resulting in a slower approach (lower slope) to the final value of total delivered traffic volume. Nevertheless, the final value differs little from that for the case of infinite transceiver and frequency resources, as we observed for the case of omnidirectional antennas.

Table 1 shows the total delivered traffic volume (corresponding to the final values of the curves in Figs. 11 and 12) for both RB-MIP and D-MIP, four values of $\theta_{\min}$, and three (F, T) pairs. This table shows that the impact of finite resources on total delivered traffic volume is small, even though the rate at which traffic is delivered (the slope of the curve) is reduced significantly.

Fig. 13 shows the impact of a higher value of the propagation constant, namely $\alpha = 4$, for unlimited transceiver and frequency resources. It is appropriate to consider higher values of α such as this because the propagation loss rate increases as the transmitting antenna is moved closer to ground level and as the density of foliage increases. The curves are qualitatively similar to those of Fig. 11, in that the curves can again be closely approximated by straight lines until the final value is reached, again with slope that is independent of $\theta_{\min}$. As in the case for $\alpha = 2$, there is a

TABLE 1
Total Delivered Traffic Volume; $\alpha = 2$

	$\theta_{min} = 360$	$\theta_{min} = 90$		$\theta_{min} = 60$		$\theta_{min} = 30$	
(F, T)		RB-MIP	D-MIP	RB-MIP	D-MIP	RB-MIP	D-MIP
(∞, ∞)	19,447	64,396	68,732	79,287	95,541	103,518	146,821
(8, 4)	20,364	64,805	68,945	79,118	95,789	102,997	146,761
(4, 2)	21,046	64,380	71,169	78,522	96,129	100,625	150,347

strong dependence of the final value of delivered traffic volume on $\theta_{\min}$. However, there is significantly less difference between the results for D-MIP and RB-MIP. For example, for $\theta_{\min} = 30$, this difference is only 16.97 percent, as compared to 41.83 percent for the corresponding case of $\alpha = 2$ and unlimited resources.

The impact of increasing α can be explained qualitatively by recognizing that, as α increases, the penalty for using longer links increases. Thus, trees typically consist of a larger number of shorter links. Consequently, under omnidirectinal MIP (and, hence, RB-MIP as well) nodes will tend to have a smaller number of downstream neighbors and, hence, the average beamwidth used for RB-MIP will tend to be less than that for smaller values of α. Therefore, there will be less of a difference in the trees produced by RB-MIP and D-MIP (which tends to produce trees in which a node's downstream neighbors are located within relatively narrow beams).

When comparing results for $\alpha = 2$ and $\alpha = 4$, we do not attach any significance to the difference in total delivered traffic volume or to network lifetime. It is not possible to compare directly the results for different values of α because we use arbitrary units of distance, and because different constant factors may be associated with propagation loss for different values of α in realistic environments.

7.3 Traffic Volume per Unit Energy

It is also of interest to study the dependence of traffic volume on $\theta_{\min}$. Fig. 14 shows $B_{X,E}$ the total number of bits delivered per unit energy over the entire lifetime of the network (in this case until no pair of live nodes is within communication range), as a function of $\theta_{\min}$, for both RB-MIP and D-MIP.

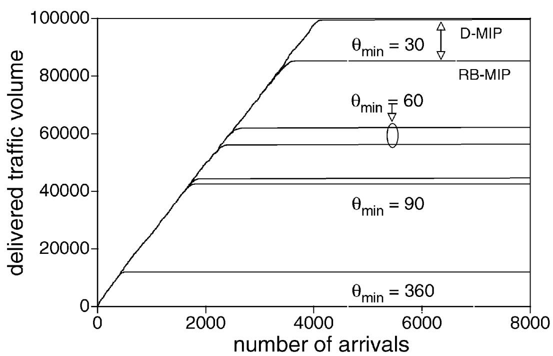


Fig. 13. Evolution of cumulative bit volume under MIP with directional antennas for D-MIP and RB-MIP; $T = \infty$, $F = \infty$, $\alpha = 4(\beta = 1; (p^T, p^R) = (0, 0))$.

Fig. 14a shows $B_{X,E}$ for $(p^T, p^R) = (0, 0)$ and $\beta = 1$. Consistent with the results presented above, D-BIP provides better performance than RB-BIP, and this difference increases as $\theta_{\min}$ decreases. There is little difference in performance for $\theta_{\min} > 90$. However, there is approximately an order of magnitude difference for $\theta_{\min} = 1$ (the smallest value for which results were obtained).

Figs. 14b and 14c show the impact of processing power, for the cases of $(p^T, p^R) = (0.01, 0.1)$ and (0.1, 1), respectively. The most obvious impact of processing power is the reduced value of $B_{X,E}$. Since energy is now expended for signal processing, less is available for RF transmission. Therefore, the overall delivered traffic volume is reduced greatly (note that the vertical scale is logarithmic in Fig. 14a and linear in the others). Moreover, the advantage of using D-BIP decreases as processing power increases, again because a smaller fraction of energy is available for RF transmission.

8 Conclusions

In this paper, we have studied the impact of the use of directional antennas on the design and performance of algorithms for energy-aware, source-initiated, session-based broadcasting and multicasting in all-wireless networks. Our primary focus has been on operation in energy-limited environments, in which the nodes are subject to hard constraints on available energy. By contrast, in applications where energy efficiency is of primary concern, it is assumed that sufficient energy is available to support all communication requirements, but its use incurs a cost.

The main contribution of this paper is the development of four tree-construction algorithms that are suitable for use with directional antennas, two for broadcasting and two for multicasting. One of the broadcast algorithms, Reduced-Beamwidth BIP (RB-BIP), uses the trees formed by BIP under the assumption of omnidirectional antennas and then reduces the beamwidth to concentrate the RF energy in the cone where it is needed. The other, Directional-BIP (D-BIP), exploits the directionality of the antennas throughout the tree-construction process. The corresponding multicasting algorithms are RB-MIP and D-MIP. The Reduced-Beam-

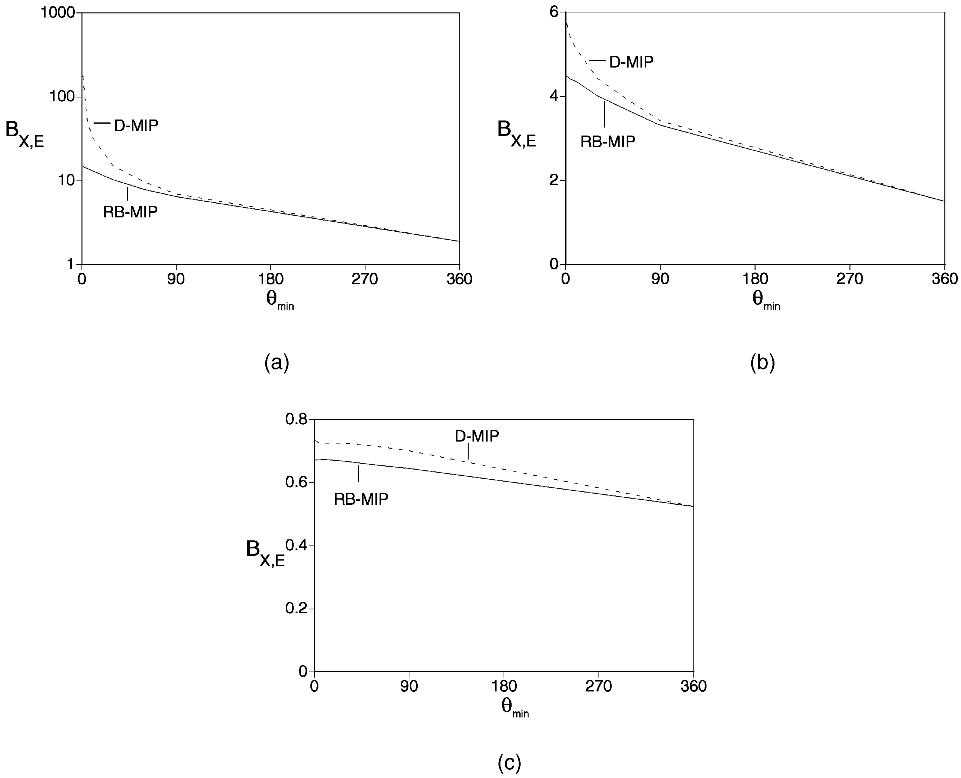


Fig. 14. Bit volume per unit energy vs. $\theta_{\min}$ for D-MIP and RB-MIP ($\beta = 1$). (a) $(p^T, p^R) = (0, 0)$, (b) $(p^T, p^R) = (0.01, 0.1)$, and (c) $(p^T, p^R) = (0.1, 1)$.

width approach can be used with any tree construction algorithm as a postalgorithm step; however, the Directional approach, which uses the incremental power philosophy we developed in our earlier research and which is characterized by higher complexity, is integrated into each step of the BIP algorithm.

We have shown that the incorporation of residual energy into local cost metrics, which results in load balancing that spreads the burden of energy use among more of the nodes, has a considerable impact on network performance. Most importantly, we have shown that the time of the first node's death can be delayed significantly, thus permitting operation at maximum throughput rates much longer than is possible when a criterion of minimum-power trees is used. Additionally, the overall volume of data that is delivered is increased. System operation is highly robust with respect to the residual-energy parameter β; values between 0.5 and 2 have been shown to work well.

The tree construction algorithms developed in this paper for directional antennas provide significant improvement in terms of network lifetime and total delivered traffic volume, as compared to their omnidirectional counterparts. The improvement is greatest for small values of $\theta_{\min}$ and low to moderate values of signal-processing power. Moreover, the "directional" versions provide significantly better performance than the "reduced-beam" versions, especially for small values of $\theta_{\min}$ and small values of processing power. However, they do so at the cost of increased complexity. The relative advantage of the directional versions decreases as the propagation constant α increases. Thus, the simpler reduced-beam versions may provide acceptable performance in many practical applications.

Our algorithms are heuristics, and no claim for optimality is made. Nevertheless, by illustrating the degree of improvement that can be obtained by exploiting the properties of directional antennas into tree-construction algorithms, we have demonstrated the benefit of the use of directional antennas in energy-constrained environments.

Acknowledgments

This work was supported by the US Office of Naval Research.

References

[1] J.E. Wieselthier, G.D. Nguyen, and A. Ephremides, "On the Construction of Energy-Efficient Broadcast and Multicast Trees in Wireless Networks," *Proc. IEEE INFOCOM 2000,* pp. 585-594, Mar. 2000.

[2] J.E. Wieselthier, G.D. Nguyen, and A. Ephremides, "Energy-Efficient Broadcast and Multicast Trees in Wireless Networks," *Mobile Networks and Applications (MONET),* vol. 7, no. 6, pp. 481-492, Dec. 2002.

[3] P.-J. Wan, G. Calinsecu, X.-Y. Li, and O. Frieder, "Minimum-Energy Broadcast Routing in Static Ad Hoc Wireless Networks," *Proc. IEEE INFOCOM 2001*, pp. 1162-1171, Apr. 2001.

[4] A. Ahluwalia, E. Modiano, and L. Shu, "On the Complexity and Distributed Construction of Energy-Efficient Broadcast Trees in Static Ad Hoc Wireless Networks," *Proc. 36th Ann. Conf. Information Sciences and Systems (CISS 2002),* Mar. 2002.

[5] F. Li and I. Nikolaidis, "On Minimum-Energy Broadcasting in All-Wireless Networks," *Proc. 26th Ann. IEEE Conf. Local Computer Networks,* pp. 193-202, 2001.

[6] G.D. Nguyen, "Construction of Broadcast and Multicast Trees with Arbitrary Cost Functions," *Proc. 36th Ann. Conf. Information Sciences and Systems (CISS 2002)* Mar. 2002.

[7] A.E.F. Clementi, P. Crescenzi, P. Penna, and P.V.G. Rossi, "On the Complexity of Computing Minimum Energy Consumption Broadcast Subgraphs," *Proc. 18th Ann. Symp. Theoretical Aspects of Computer Science,* pp. 121-131, 2001.

[8] J.E. Wieselthier, G.D. Nguyen, and A. Ephremides, "Resource Management in Energy-Limited, Bandwidth-Limited, Transceiver Limited Wireless Networks for Session-Based Multicasting," *Computer Networks,* vol. 39, no. 2, pp. 113-131, June 2002.

[9] J.E. Wieselthier, G.D. Nguyen, and A. Ephremides, "The Impact of Directional Antennas on Energy-Aware, Session-Based Multicasting in Ad Hoc Wireless Networks," *Proc. IEEE MILCOM 2001,* Oct. 2001.

[10] J.E. Wieselthier, G.D. Nguyen, and A. Ephremides, "Energy-Limited Wireless Networking with Directional Antennas: The Case of Session-Based Multicasting," *Proc. IEEE INFOCOM 2002,* pp. 190-199, June 2002.

[11] K. Wang, C.F. Chiasserini, J.G. Proakis, and R.R. Rao, "RISE: Reducing Interference and Saving Energy through Multicasting in Ad Hoc Wireless Networks," *Proc. IEEE MILCOM 2002,* Oct. 2002.

[12] A. Spyropoulos and C.S. Raghavendra, "Energy Efficient Communications in Ad Hoc Networks Using Directional Antennas," *Proc. IEEE INFOCOM 2002,* pp. 220-228, June 2002.

[13] J.E. Wieselthier, G.D. Nguyen, and A. Ephremides, "Energy-Efficient Multicasting of Session Traffic in Bandwidth- and Transceiver-Limited Wireless Networks," *J. Cluster Computing,* vol. 5, no. 2, pp. 179-192, Apr. 2002.

[14] J.E. Wieselthier, G.D. Nguyen, and A. Ephremides, "Algorithms for Energy-Efficient Multicasting in Static Ad Hoc Wireless Networks," *Mobile Networks and Applications (MONET),* vol. 6, no. 3, pp. 251-263, June 2001.

[15] M. Zorzi and R.R. Rao, "Error Control and Energy Consumption in Communications for Nomadic Computing," *IEEE Trans. Computers,* vol. 46, no. 3, pp. 279-289, Mar. 1997.

[16] D.E. Knuth, *The Art of Computer Programming—Vol. 3: Sorting and Searching,* second ed. Reading, Mass.: Addison-Wesley, 1998.

[17] J.-H. Chang and L. Tassiulas, "Energy Conserving Routing in Wireless Ad-hoc Networks," *Proc. IEEE INFOCOM 2000,* pp. 22-31, Mar. 2000.

Jeffrey E. Wieselthier received the SB degree from the Massachusetts Institute of Technology, Cambridge, in 1969, the MS degree from the Johns Hopkins University, Baltimore, Maryland, in 1971, and the PhD degree from the University of Maryland, College Park, in 1979, all in electrical engineering. He was employed at the Naval Surface Warfare Center, White Oak, Silver Spring, Maryland, from 1969 to 1979. Since 1979, he has been with the Information Technology Division of the Naval Research Laboratory, Washington, DC. He was technical program cochair of the Third IEEE Symposium on Computers and Communications in Athens, Greece, in 1998 and treasurer of the 1991 IEEE International Symposium on Information Theory in Budapest, Hungary. He won the IEEE Fred W. Ellersick Award for the best unclassified paper at MILCOM 2000. He has studied a variety of communication networking problems, including multiple access and routing in spread-spectrum networks and the use of neural networks and other approaches for network performance evaluation, optimization, and control. His current interests include wireless communication networks, with an emphasis on issues relating to energy-efficient operation. Dr. Wieselthier is a member of Eta Kappa Nu and Sigma Xi, and is a senior member of IEEE.

Gam D. Nguyen received the PhD degree in electrical engineering from the University of Maryland, College Park, Maryland, in 1990. He has been at the Naval Research Laboratory, Washington, DC, since 1991. He won the IEEE Fred W. Ellersick Award for the best unclassified paper at MILCOM 2000. His research interests include communication systems, computer communication networks, and information processing.

Anthony Ephremides received the BS degree from the National Technical University of Athens (1967), and the MS (1969) and PhD (1971) degrees from Princeton University, all in electrical engineering. He has been at the University of Maryland since 1971, and currently holds a joint appointment as professor in the Electrical and Computer Engineering Department and the Institute for Systems Research (ISR). He is cofounder of the NASA Center for Commercial Development of Space on Hybrid and Satellite Communications Networks established in 1991 at Maryland as an off-shoot of the ISR. He was a visiting professor in 1978 at the National Technical University in Athens, Greece, and in 1979 at the EECS Department of the University of California, Berkeley, and at INRIA, France. During 1985 to 1986, he was on leave at MIT and ETH in Zurich, Switzerland. He was the general chairman of the 1986 IEEE Conference on Decision and Control in Athens, Greece, and of the 1991 IEEE International Symposium on Information Theory in Budapest, Hungary. He was the technical program cochair of the IEEE INFOCOM in New York City in 1999 and of the IEEE International Symposium on Information theory in Sorrento, Italy, in 2000. He has also been the director of the Fairchild Scholars and Doctoral Fellows Program, an academic and research partnership program in Satellite Communications between Fairchild Industries and the University of Maryland. He has received several awards, including the IEEE Donald E. Fink Prize Paper Award (1992) and the Sigmobile Award of the ACM for contributions to wireless communications (1997). He has been the president of the Information Theory Society of the IEEE (1987) and has served on the Board of the IEEE (1989 and 1990). Dr. Ephremides' interests are in the areas of communication theory, communication systems and networks, queueing systems, signal processing, and satellite communications. He is a fellow of IEEE.